The Drake equation and EIDs

Jesse Brunner

2020-11-18

One of the beautiful things about mathematical equations is that they provide a way to frame our understanding and ignorance about even seemingly unknowable problems. For instance, in the late 1950's physicists and astronomers began to consider the possibility of detecting and even trying to communicate with alien life on other planets orbiting other stars. Frank Drake, an American astronomer and astrophysicist, was one of those people. In April through July 1960 he listened to radio signals from three sun-like stars for six hours a day for anything that might be recognizable as an alien signal[1]. Needless to say, he did not find any signals, but the search, along with an upcoming conference he was hosting on the search for extraterrestrial intelligence, led him to devise what became known as the "Drake equation."[2]

[1] Essentially any non-natural signals near the wavelength emitted by neutral hydrogen, which seemed like a natural place for an alien to transmit if they were trying to communicate

[2] Importantly, to me at least, he did not name the equation after himself. It was just *this equation I devised to organize my thinking*. Naming an equation or hypothesis after oneself is, in my opinion, rather poor form.

The equation was formulated as a simple way to estimate the number of civilizations in our galaxy with which communication might be possible.

$$N = R_* \times f_p \times n_e \times f_l \times f_i \times f_c \times L$$

It is, in essence, a series of rates and probabilities, the product of which give us this number, N. Working from left to right we have:

- R_* = the rate of star formation in our galaxy[3]
- f_p = the fraction of those stars that have planets.
- n_e = the average number of planets that can *potentially* support life around those stars that have planets.
- f_l = the fraction of those planets that can support life that actually develop life
- f_i = the fraction of those planets that develop life that lead to *intelligent* life (i.e., civilization)
- f_c = the fraction of those civilizations that release detectable signals into space[4]
- L = the length of time those civilizations release detectable signals into space[5].

[3] This rate is multiplied by a duration, L, so it ends up producing a number of stars.

[4] E.g., radio waves or maybe lasers.

[5] The thinking at the time was, "before they blow themselves up with nuclear weapons!" But it could also mean "until they moved from transmitting TV signals over the air to using cables or something else that did not emit signals" or "until they stopped trying to communicate with others."

Some of these numbers can be reasonably estimated. For instance, NASA and ESA have estimates that suggest $R_* = 1.5 - 3$ stars per year in our galaxy. And $f_p \times n_e$ was originally thought to be around 0.2 to 2.5, better estimates derived from the Kepler space telescope suggest $f_p \times n_e = 0.4$. Other numbers, however, are essentially a guess. How often does life arise on a potentially suitable planet? So far we have only our own planet we know of for certain. Life seems to have arisen rapidly after the initial bombardment of the planet eased, so that suggests life can arise quickly. But we have not yet[6] found life on Mars, which seems like it could be or have been suitable for life. So…who knows?

[6] As of this writing.

Some have argued that the Drake equation amounts to a Rorschach test, essentially revealing a person's existing beliefs about the likelihood of intelligent life. Indeed, individual estimates range from essentially zero to numbers in the millions or above! Others have argued (at least on YouTube) that the equation is useless because actually multiplying through these values to estimate N does nothing but reveal the consequences of our assumptions; that is, we've learned nothing new. But I disagree.

I think understanding the consequences of our assumptions or estimates is important! I also think that this equation, as it was originally intended, helps guide our thinking about the key steps or bits of knowledge we lack. If we want to be able to estimate N, we really need a better sense of those last four terms! Finally, we can turn this exercise on its head and ask, for instance, how small must $f_l \times f_i$ be to ensure that we are and have been the only intelligent life to arise in our galaxy or the whole universe[7] (setting aside the communication part).

[7] Frank & Sullivan. 2016. A New Empirical Constraint on the Prevalence of Technological Species in the Universe. Astrobiology 6:359-362. Short version: very, very unlikely!

What would a Drake equation look like for disease emergence?

In any case, this sort of thinking has been applied to a number of contexts including, recently, COVID-19 transmission[8], finding a boy/girlfriend, and, um, flakes:

[8] Mittal et al. 2020. A mathematical framework for estimating risk of airborne transmission of COVID-19 with application to face mask use and social distancing. Phys. Fluids 32, 101903

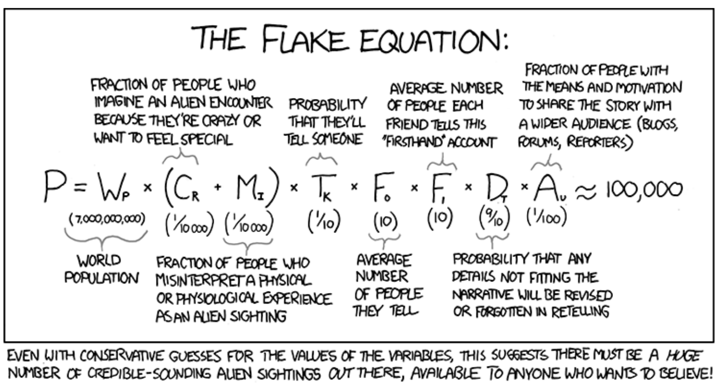

Figure 1: From https://xkcd.com/718/.

As a disease ecologist, I see a lot of similarities between the question of how many alien civilizations might be out there trying to communicate and the question of how many new diseases might emerge in any given year. I think this is a great way to think through the processes involved in emergence and their probabilities. So, in class on Thursday, this is my question for you:

WHAT WOULD A DRAKE EQUATION look like for emerging infectious diseases (EIDs)? What should go into it? What are the steps?